
Subject: Re: strange results from webget
Posted by [wlandsman](#) on Thu, 20 Dec 2012 04:22:45 GMT
[View Forum Message](#) <> [Reply to Message](#)

The quick but unsatisfying answer is that webget() doesn't seem to properly handle the "chunked" transfer encoding used by your Website. (This hasn't been an issue for most of the astronomy websites that webget() is used for.) Some options for you are

(1) I am pretty sure that Dominic Zarro's http code in the Solarsoft library does handle chunked transfer

<http://sohowww.nascom.nasa.gov/solarsoft/gen/idl/http/>

but I haven't used it.

(2) I believe the IDLNetURL object will handle chunked transfer encoding but again I haven't used it.

(3) Update webget() so that it can handle chunked encoding.

--Wayne

On Wednesday, December 19, 2012 5:10:28 PM UTC-5, JP wrote:

> G'day all,

>

>

>

> I am using the webget function from astrolib <http://idlastro.gsfc.nasa.gov/> and i am finding some weird things.

>

> This is my example:

>

> If I call webget to retrieve the information from an url like this:

>

> [http://thredds0.nci.org.au/thredds/dodsC/u39/modis/lpdaac-mosaics-cmar/v1-hdf4/aust/MOD09A1.005/2004.12.10/MOD09A1.2004.345.aust.005.b01.500m_0620_0670nm_refl.hdf.gz.ascii?500m_0620_0670nm_refl\[5109:1:5343\]\[7451:1:7472\]](http://thredds0.nci.org.au/thredds/dodsC/u39/modis/lpdaac-mosaics-cmar/v1-hdf4/aust/MOD09A1.005/2004.12.10/MOD09A1.2004.345.aust.005.b01.500m_0620_0670nm_refl.hdf.gz.ascii?500m_0620_0670nm_refl[5109:1:5343][7451:1:7472])

>

> If you try that in a browser you'll see it returns meaningful data. (for thos interested that's a small subset of a MODIS reflectance product for Australia).

>

>

>

> I try that with webget doing:

>

>

>

```

> url = ' http://thredds0.nci.org.au/thredds/dodsC/u39/modis/lpdaac-mo
saics-cmar/v1-hdf4/aust/MOD09A1.005/2004.12.10/MOD09A1.2004.
345.aust.005.b01.500m_0620_0670nm_refl.hdf.gz.ascii?500m_062
0_0670nm_refl[5109:1:5343][7451:1:7472]'
>
>
>
> a = webget(url)
>
>
>
> so far so good.
>
> However, when i explore the contents of a.text, in a few cases the result doesnt match what
you get in your browser. For example:
>
>
>
> for i=68, 71 do print, a.text[i]
>
>
>
> I get:
>
>
>
> [56], 1921, 1980, 1785, 1954, 2078, 2072, 1822, 1822, 1401, 1494, 1494, 1840
>
> 2000
>
> , 2218, 2050, 2050, 1798, 1719, 1683, 1808, 1652, 1940, 1703
>
> [57], 1945, 1980, 1921, 1927, 1865, 1865, 1944, 1678, 1290, 1255, 1477, 1901, 1840, 1996,
1916, 1631, 1762, 1756, 1756, 1652, 1652, 1641
>
>
>
> that's different from what i expect , looking at the result in my browser which looks:
>
>
>
> [56], 1921, 1980, 1785, 1954, 2078, 2072, 1822, 1822, 1401, 1494, 1494, 1840, 2218, 2050,
2050, 1798, 1719, 1683, 1808, 1652, 1940, 1703
>
> [57], 1945, 1980, 1921, 1927, 1865, 1865, 1944, 1678, 1290, 1255, 1477, 1901, 1840, 1996,
1916, 1631, 1762, 1756, 1756, 1652, 1652, 1641
>
>

```

>
> the problem is repeated a few other times.
>
>
>
> Any ideas? I know wget uses SOCKET and other things like that, I didn't go that far in
exploring where the problem could be in.
>
>
>
> Thanks
>
>
>
> JP
